THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 52, No. 1, January, 1973
Printed in U.S.A.

Batch Input to a Multiserver Queue with Constant Service Times

By A. KUCZURA

(Manuscript received July 26, 1972)

Delay probability formulas for batch input to a finite number of constant-holding-time servers are derived under the assumption of statistical equilibrium. The service-delay distribution (delay until a first request from the batch enters service) is given in terms of the roots of a transcendental equation, while the probability of no service-delay and the average delay are expressed directly in terms of the number of servers, the holding time, and the parameters of the input process. A numerical example with a fixed batch size is discussed.

I. INTRODUCTION

Batch arrivals constitute an important class of input processes in the theory of queues. The investigation of the problem of batcb input to a group of constant-holding-time servers was motivated by the existence of installations witb multiple Automatic Calling Units (ACU). Customer-based computer equipment controlling tbe ACU's is capable of originating simultaneous requests. The dial-tone markers, the first common control equipment in a No. 5 Crossbar central office to serve the requests, can be modeled as a group of constant-holding-time servers.

Another example comes from an information transmission system. Messages containing a (small) random number of characters (a batch of characters) arrive according to a Poisson process and must be transmitted to some destination. Delayed messages are stored in a buffer. Since tbe transmission time per character is usually fixed, this system provides anotber example of the model studied.

In Section II, the matbematical model used in this study is described and the input process defined; the state equations are written and used to derive the generating function for the equilibrium state probabilities. The probability of no service-delay is found in Section

III, while the average delay is computed in Section IV. The service-delay distribution is given in terms of the roots of a transcendental equation in Section V. A numerical example with a fixed batch size is discussed in Section VI. The effect of this batching scheme on the average delay and the service-delay probability is examined.

II. MATHEMATICAL MODEL

The model studied here is that of a queuing system with a finite number of servers, batch arrivals, and constant holding time. The assumptions are

(*i*) Requests arrive according to a compound Poisson process, that is, requests arrive in groups or batches and the instants at which the batches arrive constitute a Poisson process.
(*ii*) There are c servers and each request has access to any one of them.
(*iii*) All requests have the same constant service time, τ.
(*iv*) The delayed batches wait until service becomes available and are served in order of arrival. The service discipline for requests within a batch is arbitrary, i.e., not specified here.
(*v*) The system is in statistical equilibrium.

Systems with simple Poisson input have been studied as early as 1920, when A. K. Erlang[1] obtained expressions for the probability of delay for arbitrary values of c and the average delay for $c = 1, 2,$ and 3. The first complete treatment of such systems was by Pollaczek.[2,3] Crommelin,[4,5] using a method which is simpler than that of Erlang or Pollaczek, also derived general formulas for the probability of delay, the average delay, and the delay distribution. A simplified and concise account of Crommelin's work is given by A. Descloux,[6] who also shows how Pollaczek's formulas can be deduced from Crommelin's results. The development herein is an extension of Crommelin's results to the case of compound Poisson input using the simpler methods employed by Descloux.

We now define the input process. Consider events which happen in groups rather than singly, that is, requests arriving in batches at a group of c servers. For $k = 1, 2, \cdots$, let $N_k(t)$ be a Poisson process with intensity λ_k which governs the arrival of k-sized batches. Assume independence of the processes $N_k(t)$, $k = 1, 2, \cdots$. Let $N(t)$ be the total number of requests that have arrived in the interval $(0, t]$.

Then

$$N(t) = \sum_{k=1}^{\infty} kN_k(t) \tag{1}$$

is called a *compound* Poisson process (Ref. 7, page 271).

The probability that an arriving batch is of size k is equal to λ_k/λ, where

$$\lambda = \sum_{k=1}^{\infty} \lambda_k.$$

From eq. (1), we see that the mean and variance of the number of arrivals per unit time are

$$\mu_1 = \sum_{k=1}^{\infty} k\lambda_k \qquad \text{and} \qquad \mu_2 = \sum_{k=1}^{\infty} k^2\lambda_k, \tag{2}$$

respectively. The generating function of the probability distribution $\pi_n(t) = P\{N(t) = n\}$, $n = 0, 1, 2, \cdots$, is given by

$$\pi(t,z) = \sum_{n=0}^{\infty} \pi_n(t)z^n = e^{t\beta(z)},$$

where

$$\beta(z) = \sum_{n=1}^{\infty} \lambda_n z^n - \lambda,$$

and hence the probabilities $\pi_n(t)$ are given by

$$\pi_n(t) = e^{-\lambda t} \sum_{\mathcal{S}_n} \frac{(\lambda_1 t)^{k_1}(\lambda_2 t)^{k_2} \cdots (\lambda_n t)^{k_n}}{k_1!k_2!\cdots k_n!}, \qquad n = 0, 1, 2, \cdots, \tag{3}$$

where $\mathcal{S}_n$ is the class of all sets of nonnegative integers $\{k_1, k_2, \cdots, k_n\}$ such that $k_1 + 2k_2 + \cdots + nk_n = n$.

The expression given in (3) is not suited for computing $\pi_n(t)$. These probabilities are more conveniently computed from the recurrence relation

$$\begin{aligned} \pi_{k+1}(t) &= \frac{t}{k+1} \sum_{j=0}^{k} (k - j + 1)\lambda_{k-j+1}\pi_j(t), \qquad k = 0, 1, 2, \cdots, \\ \pi_0(t) &= e^{-\lambda t}. \end{aligned} \tag{4}$$

Equation (4) is easily obtained from the relation $k!\pi_k(t) = \pi^{(k)}(t, 0)$, where the superscript denotes differentiation with respect to z.

Special cases of the compound Poisson process are obtained by choosing different convergent sequences of the positive constants

$\lambda_1, \lambda_2, \cdots$. One such sequence is obtained by setting $\lambda_j = \sigma/r^{j-1}$, $r > 1$, $j = 1, 2, \cdots$. This special example has become known as the "stuttering" Poisson process.[8] In this case, a simple expression for $\pi_n(t)$ can be obtained by noting that the generating function has a power series expansion in z, the coefficients of which are given in terms of the Laguerre polynomials L_n, that is,

$$e^{t\beta(z)} = e^{-\lambda t}\left(1 - \frac{z}{r}\right) \sum_{n=0}^{\infty} L_n(-\sigma t r) \left(\frac{z}{r}\right)^n,$$

since

$$\beta(z) = \frac{\sigma z}{1 - z/r} - \lambda.$$

It follows that

$$\begin{aligned} \pi_0(t) &= e^{-\lambda t} \\ \pi_n(t) &= \frac{e^{-\lambda t}}{r^n}\left[L_n(-\sigma t r) - L_{n-1}(-\sigma t r)\right], \qquad n = 1, 2, \cdots. \end{aligned} \tag{5}$$

We will now obtain the equilibrium state equations, and find the probability generating function of the stationary distribution for the general case. Let $X(t)$ be the number of requests (waiting or in service) in the system at time t. Let

$$p_{ij}(t) = P\{X(t) = j \mid X(0) = i\}$$

be the transition probability functions of the process $\{X(t),\ t \geqq 0\}$. It is clear that $X(t)$ is not Markovian. If, however, we examine

$$X_k = X(k\tau), \qquad k = 0, 1, 2, \cdots, \tag{6}$$

we see that this sequence is Markovian and, in fact, $\{X_k, k=0, 1, 2, \cdots\}$ is a homogeneous Markov chain with one-step transition probabilities

$$p_{ij} = P\{X_{k+1} = j \mid X_k = i\}, \qquad k = 0, 1, 2, \cdots,$$

given by

$$p_{ij} = \begin{cases} \pi_j(\tau), & \text{for } 0 \leqq i \leqq c \\ \pi_{j-i+c}(\tau), & \text{for } c < i \leqq j + c \\ 0, & \text{for } j + c < i. \end{cases}$$

We will be interested in the distribution of the number of requests in the system encountered by an arbitrarily arriving batch (a batch arriving at a time point a long way from the origin, i.e., after statistical equilibrium has been reached). But since the instants at which the batches arrive constitute a Poisson process, this distribution is the

same as the stationary distribution

$$p_j = \lim_{t \to \infty} p_{ij}(t), \qquad j = 0, 1, 2, \cdots,$$

of the process $\{X(t), t \geqq 0\}$. Moreover, if this limit exists, then so does

$$\lim_{k \to \infty} P\{X_k = j \mid X_0 = i\}$$

and they are equal. Consequently, the distribution of interest to us is given by the stationary distribution of the imbedded Markov chain (6). This distribution is obtained by solving the Chapman-Kolmogorov equations

$$\begin{aligned} p_0 &= \pi_0(\tau) a_c \\ p_n &= \pi_n(\tau) a_c + \sum_{m=c+1}^{n+c} p_m \pi_{n-m+c}(\tau), \qquad n = 1, 2, \cdots, \end{aligned} \tag{7}$$

where

$$a_n = \sum_{m=0}^{n} p_m.$$

We will assume that $\mu_1 \tau < c$ so that the stationary distribution $\{p_j\}$ exists.

We need to solve the system (7) for the unknowns p_n. To do this we introduce the probability generating function

$$f(z) = \sum_{n=0}^{\infty} p_n z^n.$$

Multiplying both sides of (7) by z^n and summing over n, we have

$$\begin{aligned} f(z) &= a_c e^{\tau\beta(z)} + \sum_{n=1}^{\infty} \sum_{j=1}^{n} p_{c+j} \pi_{n-j}(\tau) z^n \\ &= a_c e^{\tau\beta(z)} + \sum_{j=1}^{\infty} \sum_{n=j}^{\infty} p_{c+j} \pi_{n-j}(\tau) z^n \\ &= a_c e^{\tau\beta(z)} + \frac{1}{z^c} e^{\tau\beta(z)} [f(z) - g(z)] \end{aligned}$$

where

$$g(z) = \sum_{n=0}^{c} p_n z^n.$$

Thus the probability generating function of the sequence p_n,

$n = 0, 1, \cdots$, is given hy

$$f(z) = \frac{g(z) - z^c a_c}{1 - z^c e^{-\tau\beta(z)}}. \tag{8}$$

III. PROBABILITY OF NO SERVICE-DELAY

We say that the service of a hatch is delayed if upon its arrival all servers are husy. Hence, the prohahility of no service-delay will he defined hy a_{c-1}, that is, the prohahility of at most $c - 1$ servers husy. An explicit expression for this prohahility will now he found.

We start with eq. (8). Since the p_n's are prohahilities, $f(z)$ is holomorphic in $|z| \leqq 1$ and, therefore, the zeros of the denominator and numerator in $|z| \leqq 1$ must he the same and have the same multiplicities. We will show that the denominator of (8) has c distinct roots in $|z| \leqq 1$ and that all of them, with the exception of $z = 1$, lie inside the unit circle.

For $|z| = 1 + \delta$, with δ sufficiently small and positive, we have

$$|e^{\tau\beta(z)}| \leqq e^{-\lambda\tau} \exp \tau \left\{ \sum_{k=1}^{\infty} \lambda_k |z|^k \right\} = e^{\tau\mu_1\delta + 0(\delta^2)}$$

where μ_1 is defined hy (2).

Since $\tau\mu_1 < c$ hy assumption, we have

$$e^{\tau\mu_1\delta+0(\delta^2)} < (1 + \delta)^c = |z|^c,$$

and hy Rouché's theorem, the equation

$$e^{\tau\beta(z)} - z^c = 0$$

has exactly c roots within the region $|z| = 1 + \delta$. Let these roots be denoted hy $z_1, z_2, \cdots, z_{c-1}, z_c(= 1)$. Then

(*i*) $1, z_1, z_2, \cdots, z_{c-1}$ are distinct.
(*ii*) $|z_n| < 1$ for $n = 1, 2, \cdots, c - 1$.

To prove (*i*), first note that the root $z = 1$ is simple because

$$\lim_{z \to 1} \frac{1 - z^c e^{-\tau\beta(z)}}{z - 1} = \tau\mu_1 - c \neq 0\cdot$$

Similarly, for any root z_i, $i = 1, 2, \cdots, c - 1$,

$$\lim_{z \to z_i} \frac{1 - z^c e^{-\tau\beta(z)}}{z - z_i} = z_i^{c-1} e^{-\tau\beta(z_i)} \left[\tau \sum_{k=1}^{\infty} k\lambda_k z_i^k - c \right]\cdot$$

The first two factors cannot vanish for any admissible choice of the root z_i, so that if z_i is to be a root of second or higher order, we must have

$$\tau(\lambda_1 z_1 + 2\lambda_2 z_2^2 + 3\lambda_3 z_3^3 + \cdots) = c.$$

But this is not possible since

$$\tau|\lambda_1 z_1 + 2\lambda_2 z_2^2 + 3\lambda_3 z_3^3 + \cdots| \leqq \tau\mu_1 < c,$$

and the roots $1, z_1, z_2, \cdots, z_{c-1}$ are therefore all distinct.

To prove (ii), suppose that $|z_n| = 1$ for some n, $n = 1, 2, \cdots, c-1$, then $|\exp[\tau\beta(z_n)]| = 1$ and hence the real part of $\tau\beta(z_n)$ must be zero, that is, $\mathcal{R}[\beta(z_n)] = 0$. Hence we must bave

$$\mathcal{R}[\beta(z_n)] = \mathcal{R}\left\{\sum_{k=1}^{\infty} \lambda_k \left(1 - z_n^k\right)\right\} = 0\cdot$$

Since all terms in the sum are nonnegative, we have $\mathcal{R}(1 - z_n^k) = 0$ for all k, and therefore $z_n = 1$, contrary to the assumption. It follows that $|z_n| < 1$ for $n = 1, 2, \cdots, c-1$.

Since the numerator of (8) is a polynomial of degree c, $f(z)$ has the form

$$f(z) = A\frac{(z-1)(z-z_1)(z-z_2)\cdots(z-z_{c-1})}{1 - z^c e^{-\tau\beta(z)}}. \tag{9}$$

The condition $f(1) = 1$ determines A:

$$A = \frac{\mu_1\tau - c}{(1-z_1)(1-z_2)\cdots(1-z_{c-1})}. \tag{10}$$

In computing a_{c-1} it is convenient to introduce the generating function

$$F(z) = \sum_{n=0}^{\infty} a_n z^n\cdot$$

Then, since $a_n - a_{n-1} = p_n$, $n = 1, 2, \cdots$, we have

$$(1-z)F(z) = f(z),$$

or

$$F(z) = \frac{f(z)}{1-z}.$$

Now, making use of (9), we obtain

$$F(z) = -A\frac{(z-z_1)(z-z_2)\cdots(z-z_{c-1})}{1 - z^c e^{-\tau\beta(z)}}.$$

The probability of no service-delay, a_{c-1}, is given by the coefficient of z^{c-1} in the expansion of $F(z)$:

$$a_{c-1} = -A = \frac{c - \mu_1 \tau}{(1 - z_1)(1 - z_2)\cdots(1 - z_{c-1})}. \tag{11}$$

An expression for a_{c-1} which does not involve the roots z_i is obtained through an application of the generalized argument-principle (Ref. 9, page 151). That is, suppose $\psi(z)$ is holomorphic and $\phi(z)$ is meromorphic on and inside the contour C. Let α_k, $k = 1, 2, \cdots$, be the zeros with multiplicities r_k, and β_k, $k = 1, 2, \cdots$, the poles with multiplicities s_k of the function $\phi(z)$ inside C. Then the generalized argument-principle states that

$$\frac{1}{2\pi i}\int_C \psi(z)\frac{\phi'(z)}{\phi(z)}dz = \sum_k r_k\psi(\alpha_k) - \sum_k s_k\psi(\beta_k).$$

Taking the logarithm of eq. (11), we have

$$\log a_{c-1} = \log (c - \mu_1\tau) - \sum_{i=1}^{c-1} \log (1 - z_i).$$

We will eliminate the roots z_i from the second term of the right-hand side of the preceding equation. Let

$$\phi(z) = e^{\tau\beta(z)} - z^c,$$

and note that $\phi(z)$ has simple zeros at $z = z_1, z_2, \cdots, z_{c-1}$. Choose $\psi(z)$ as the principal branch of $\log (1 - z)$. The generalized argument-principle yields

$$\sum_{n=1}^{c-1} \log (1 - z_n) = \frac{1}{2\pi i}\int_C \log (1 - z)d[\log \phi(z)] = J$$

where C is the contour $|z| = 1 - \epsilon$ and $\epsilon(>0)$ is chosen so that z_n, $n = 1, 2, \cdots, c - 1$, lie inside C but $z_c(=1)$ is exterior to C. We will now show that

$$J = \log (c - \tau\mu_1) + \sum_{n=1}^{\infty} \frac{1}{n} \sum_{j=nc}^{\infty} \pi_j(n\tau),$$

and hence

$$\log a_{c-1} = -\sum_{n=1}^{\infty} \frac{1}{n} \sum_{j=nc}^{\infty} \pi_j(n\tau). \tag{12}$$

Note first that the principal branch of $\log (1 - z) \dfrac{d}{dz} [z^{c-1} \log (1 - z)]$

is holomorphic in $|z| \leqq 1 - \epsilon$. Since its integral on C is zero, we have

$$J = \frac{1}{2\pi i} \int_C \log (1 - z) d \left[\log \frac{1 - z^{-c} e^{\tau \beta(z)}}{1 - z^{-1}} \right].$$

Integration by parts yields

$$J = -\frac{1}{2\pi i} \int_C \log \left[\frac{1 - z^{-c} e^{\tau \beta(z)}}{1 - z^{-1}} \right] \frac{dz}{z - 1}.$$

The integrand above has a simple pole at $z = 1$, and its residue there is equal to $\log (c - \mu_1 \tau)$. Choose $\delta (>0)$ such that the only zeros of $\phi(z)$ in the disk $|z| \leqq 1 + \delta$ are $1, z_1, z_2, \cdots, z_{c-1}$, and let C_1 be the contour $|z| = 1 + \delta$. Noting that the integral of $\log (1 - z^{-1})/(z - 1)$ on C_1 vanishes, we have

$$J = \log (c - \mu_1 \tau) - \frac{1}{2\pi i} \int_{C_1} \log [1 - z^{-c} e^{\tau \beta(z)}] \frac{dz}{z - 1}.$$

Now since $|z^{-c} e^{\tau \beta(z)}| < 1$ on C_1, the power series for $\log [1 - z^{-c} e^{\tau \beta(z)}]$ converges uniformly on C_1, and termwise integration is allowed, so that

$$\begin{aligned} J &= \log (c - \mu_1 \tau) + \frac{1}{2\pi i} \sum_{n=1}^{\infty} \frac{1}{n} \int_{C_1} e^{\tau n \beta(z)} \frac{z^{-nc}}{z - 1} dz \\ &= \log (c - \mu_1 \tau) + \sum_{n=1}^{\infty} \frac{1}{n} \left[1 - \frac{1}{2\pi i} \sum_{j=0}^{\infty} \int_{C_1} e^{\tau n \beta(z)} z^{-nc+j} dz \right]. \end{aligned}$$

Expanding the integrand in powers of z, and integrating term by term, we see that the integral is zero for all terms except one, and that there it is equal to $2\pi i$ times the coefficient of z^{-1}. But the coefficient of z^{-1} is exactly $\pi_{nc-j-1}(n\tau)$, so that

$$\begin{aligned} J &= \log (c - \mu_1 \tau) + \sum_{n=1}^{\infty} \frac{1}{n} \left[1 - \sum_{j=0}^{nc-1} \pi_{nc-j-1}(n\tau) \right] \\ &= \log (c - \mu_1 \tau) + \sum_{n=1}^{\infty} \frac{1}{n} \sum_{j=nc}^{\infty} \pi_j(n\tau), \end{aligned}$$

and this is the result stated earlier.

For the case of a simple Poisson input, we have $\lambda_1 = \lambda$, $\lambda_j = 0$, $j = 2, 3, \cdots$, and (12) reduces to [Crommelin, Ref. 5, eq. (5)]

$$\log a_{c-1} = - \sum_{n=1}^{\infty} \frac{1}{n} \sum_{j=nc}^{\infty} \frac{(\lambda n \tau)^j}{j!} e^{-\lambda n \tau}.$$

Note that this expression and eq. (12) differ only in the terms $\pi_j(n\tau)$ and $((\lambda n\tau)^j/j!)e^{-\lambda n\tau}$, wbicb represent tbe same prohahilities in two different systems: hoth are equal to the prohahility of exactly j arrivals in the time interval $n\tau$. As we sball see later, these probahilities appear again in the expressions for the average delay and the service-delay distrihution.

For "stuttering" Poisson input, eq. (12) reduces to

$$\log a_{c-1} = -\sum_{n=1}^{\infty} \frac{e^{-[\sigma n\tau r/(r-1)]}}{n} \sum_{j=nc}^{\infty} \frac{1}{r^j} [L_j(-\sigma n\tau r) - L_{j-1}(-\sigma n\tau r)]$$

where the $L_n(\xi)$ are Laguerre polynomials.

IV. AVERAGE DELAY

Under equilibrium conditions, the average delay D is equal to tbe average amount of waiting per unit of time divided hy tbe average number of arrivals per unit of time. Tbe average amount of waiting per unit of time is equal to

$$\sum_{n=c+1}^{\infty} (n-c)p_n ,$$

so that D is given by

$$D = \frac{1}{\mu_1} \sum_{n=c+1}^{\infty} (n-c)p_n = \frac{1}{\mu_1} \sum_{n=0}^{\infty} np_n - \tau$$

where μ_1 is the average number of arrivals per unit of time defined by (2). An explicit expression for D can be readily obtained hy noting that

$$\sum_{n=0}^{\infty} np_n = \lim_{z \to 1} \left[\frac{d}{dz} f(z)\right], \qquad |z| < 1 ,$$

where $f(z)$ is the generating function given by (8). Straightforward differentiation of (8) and application of L'Hospital's rule lead to

$$\frac{D}{\tau} = \frac{1}{\mu_1\tau} \sum_{i=1}^{c-1} \frac{1}{1-z_i} + \frac{\mu_2\tau + \mu_1\tau(\mu_1\tau - 1) - c(c-1)}{2\mu_1\tau(c - \mu_1\tau)} \tag{13}$$

where $z_1, z_2, \cdots, z_{c-1}$ are the roots defined previously. Again we wish to eliminate these roots. The method used in the previous section suggests the application of the generalized argument-principle witb

$(1 - z)^{-1}$ as $\psi(z)$, and $\phi(z)$ as before. Thus we have

$$\sum_{n=1}^{c-1} \frac{1}{1 - z_n} = \frac{1}{2\pi i} \int_C \frac{1}{1 - z} \frac{\phi'(z)}{\phi(z)}\, dz = K. \tag{14}$$

Noting that the integral of $(1 - z)^{-1} \dfrac{d}{dz} \{\log [z^{c-1}(1 - z)]\}$ on C is equal to $(c - 1)2\pi i$ (the residue at the simple pole $z = 0$ is $c - 1$), eq. (14) becomes

$$K = (c - 1) + \frac{1}{2\pi i} \int_C \frac{1}{(1 - z)}\, d\left[\log \frac{1 - z^{-c}e^{\tau\beta(z)}}{1 - z^{-1}}\right].$$

Integration by parts yields

$$K = (c - 1) - \frac{1}{2\pi i} \int_C \log \left[\frac{1 - z^{-c}e^{\tau\beta(z)}}{1 - z^{-1}}\right] \frac{dz}{(1 - z)^2}. \tag{15}$$

The integrand in (15) has a pole of second order at $z = 1$, and its residue there is found to be

$$\frac{c(c - 1) - \tau\mu_1(\tau\mu_1 - 1) - \tau\mu_2}{2(c - \tau\mu_1)} - (c - 1).$$

Consequently,

$$K = -\frac{c(c - 1) - \tau\mu_1(\tau\mu_1 - 1) - \tau\mu_2}{2(c - \tau\mu_1)} - \frac{1}{2\pi i} \int_{C_1} \log \left[\frac{1 - z^{-c}e^{\tau\beta(z)}}{1 - z^{-1}}\right] \frac{dz}{(1 - z)^2}. \tag{16}$$

Combining (14) and (16) and noting that the integral of

$$(1 - z)^{-2} \log (1 - z^{-1})$$

vanishes on C_1, we obtain

$$\frac{D}{\tau} = -\frac{1}{2\pi i\mu_1} \int_{C_1} \log[1 - z^{-c}e^{\tau\beta(z)}] \frac{dz}{(1 - z)^2}.$$

Recall now that $|z^{-c}e^{\tau\beta(z)}| < 1$ on C_1, and hence $\log [1 - z^{-c}e^{\tau\beta(z)}]$ has a uniformly convergent power series representation in $z^{-c}e^{\tau\beta(z)}$ on C_1, so that

$$\frac{D}{\tau} = \frac{1}{\mu_1\tau} \frac{1}{2\pi i} \sum_{n=1}^{\infty} \frac{1}{n} \int_{C_1} \frac{e^{n\tau\beta(z)}z^{-nc}}{(1 - z)^2}\, dz. \tag{17}$$

The integrand in (17) has poles at $z = 0$ and $z = 1$ with residues

$$\sum_{k=0}^{nc-1} (nc - k)\pi_k(n\tau)$$

and

$$n(\mu_1\tau - c),$$

respectively, giving us the final result

$$\frac{D}{\tau} = \frac{1}{\mu_1\tau} \sum_{n=1}^{\infty} \frac{1}{n} \sum_{k=1}^{\infty} k\pi_{nc+k}(n\tau). \tag{18}$$

V. SERVICE-DELAY DISTRIBUTION

For the purpose of obtaining the service-delay distribution (delay until a first request from the batch enters service), let us define $g_m(t)$ as the probability that among the requests present at some time t_0, at most m of them are still in the system at time $t_0 + t$. Considering the state corresponding to $g_{mc+c-1}(t)$, we see that, at most, $mc + c - 1$ of the requests preceding the given batch will be in progress at time t later, and consequently, at most, $c - 1$ of them at time $m\tau + t$ later. This is the condition for the service-delay d to be less than $m\tau + t$, or in symbols

$$P\{d < m\tau + t\} = g_{mc+c-1}(t), \qquad 0 \leqq t < \tau. \tag{19}$$

To determine $g_m(t)$, we introduce the generating function

$$G(z,t) = \sum_{m=0}^{\infty} g_m(t)z^m.$$

Upon noting that $\sum_{m=0}^{n} \pi_{n-m}(t)g_m(t) = a_n$, we have

$$\begin{aligned} G(z,t) &= e^{-t\beta(z)} \sum_{m=0}^{\infty} g_m(t)z^m \sum_{n=0}^{\infty} \pi_n(t)z^n \\ &= e^{-t\beta(z)} \sum_{n=0}^{\infty} z^n \sum_{m=0}^{\infty} \pi_{n-m}(t)g_m(t) \\ &= e^{-t\beta(z)} \sum_{n=0}^{\infty} a_n z^n = e^{-t\beta(z)}F(z). \end{aligned}$$

Substituting for $F(z)$ we obtain

$$G(z,t) = -A\,\frac{(z - z_1)(z - z_2)\cdots(z - z_{c-1})e^{-t\beta(z)}}{1 - z^c e^{-\tau\beta(z)}}. \tag{20}$$

A direct expansion of the right-hand side of (20) in powers of z is not desirable because the coefficients of z involve sums the terms of which have alternate signs and, therefore, are not well suited for computation. To circumvent this difficulty, we first obtain a Laurent series expansion of $G(z,t)$ in the annulus $1 < |z| < |\xi|$, where ξ is that root of $z^c \exp[-\tau\beta(z)] - 1 = 0$ which has the smallest modulus exceeding 1. The existence of such a root can be proved as follows. Since $x^c \exp[-\tau\beta(x)]$ takes on the value 1 at $x = 1$ (x = real part of z), has a positive derivative there, and vanishes at infinity, the equation $z^c \exp[-\tau\beta(z)] - 1 = 0$ has at least one root outside the unit circle.

For $1 < |z| < |\xi|$, the absolute value of $z^{-c} \exp[\tau\beta(z)]$ is less than unity. Expanding the denominator in powers of $z^{-c} \exp[\tau\beta(z)]$ and the exponential function in powers of z, and collecting like-power terms, we obtain, for $1 < |z| < |\xi|$,

$$G(z,t) = \sum_{k=0}^{\infty} \left\{ \sum_{n=0}^{c-1} q_n \sum_{j=0}^{\infty} \pi_{cj+c+k-n}[(j+1)\tau - t] \right\} z^k \tag{21}$$

where q_n is the coefficient of z^n in the polynomial

$$A(z - z_1)(z - z_2) \cdots (z - z_{c-1}).$$

Since $z = 1$ is the only singularity of $G(z, t)$ in $|z| < |\xi|$ (a simple pole with residue -1), $G(z, t) + (z - 1)^{-1}$ is holomorphic in $|z| < |\xi|$ and hence for $|z| < |\xi|$

$$G(z,t) + \frac{1}{z-1} = \text{expansion (21)} + \sum_{n=1}^{\infty} z^{-n}.$$

Therefore, for $|z| < 1$, we must have

$$G(z,t) = \sum_{n=0}^{\infty} z^n + \text{expansion (21)} + \sum_{n=1}^{\infty} z^{-n}.$$

From this equation, we obtain the service-delay distribution

$$P\{d < m\tau + t\} = 1 - \sum_{n=0}^{c-1} q_n \sum_{j=0}^{\infty} \pi_{(m+j+2)c-n-1}[(j+1)\tau - t],$$
$$m = 0, 1, 2, \cdots, \qquad 0 \leqq t < \tau. \tag{22}$$

VI. A NUMERICAL EXAMPLE

We examine a fixed-size batching scheme which provides some insight into the effect of batch arrivals on the average delay and the

probability of service-delay. Suppose customers arrive in batches of size m. Then $\lambda_j = 0$ for $j \neq m$, $\mu_1 = m\lambda_m$ and

$$\pi_j(t) = \begin{cases} e^{-\lambda_m t} \dfrac{(\lambda_m t)^k}{k!}, & \text{for} \quad j = km \\ 0 \quad , & \text{otberwise.} \end{cases}$$

Figure 1 sbows tbe average delay experienced for an arbitrary customer as a function of the occupancy $\rho = \tau\mu_1/c$, for various values of m. We assume that tbe bolding time is unity, and that $c = 4$.

Equation (18) written in the form

$$D = \frac{1}{\mu_1} \sum_{n=1}^{\infty} \frac{1}{n} \left\{ n(\mu_1 - c + ce^{-n\lambda_m}) + \sum_{j=1}^{nc-1} (nc - j)\pi_j(n) \right\}$$

was used to obtain the curves drawn in Fig. 1. We might point out here tbat the above series converges slowly when the occupancy is near unity. In the interest of speedy computation it may be necessary to solve for the roots of the denominator in eq. (8) and then use (13) to calculate the average delay. The same remarks apply to eq. (12) which is used to compute the prohability of no delay.

Because holding times are constant, several interesting phenomena are observed. First, if the batch size is an integer multiple of the number of servers, say $m = kc$, then the mean time until the service of an arriving batch (or the first customer from the batch) begins is the same as the average delay in a one-server system with single Poisson arrivals and holding time k. From the Pollaczek-Khintchine formula, this number is given by

$$\frac{k\rho}{2(1 - \rho)}.$$

Hence tbe mean delay wbicb an arbitrary customer experiences is the average of the above number and tbe mean delay experienced by the last customer in the batcb to be served. Thus we have

$$D = \frac{k - 1}{2} + \frac{\rho k}{2(1 - \rho)} \qquad (m = kc).$$

Note that as $\rho \to 0$, $D \to (k - 1)/2$.

On the other hand, if the number of servers is an integer multiple of the hatch size, say $c = jm$, tben the system may be viewed as a

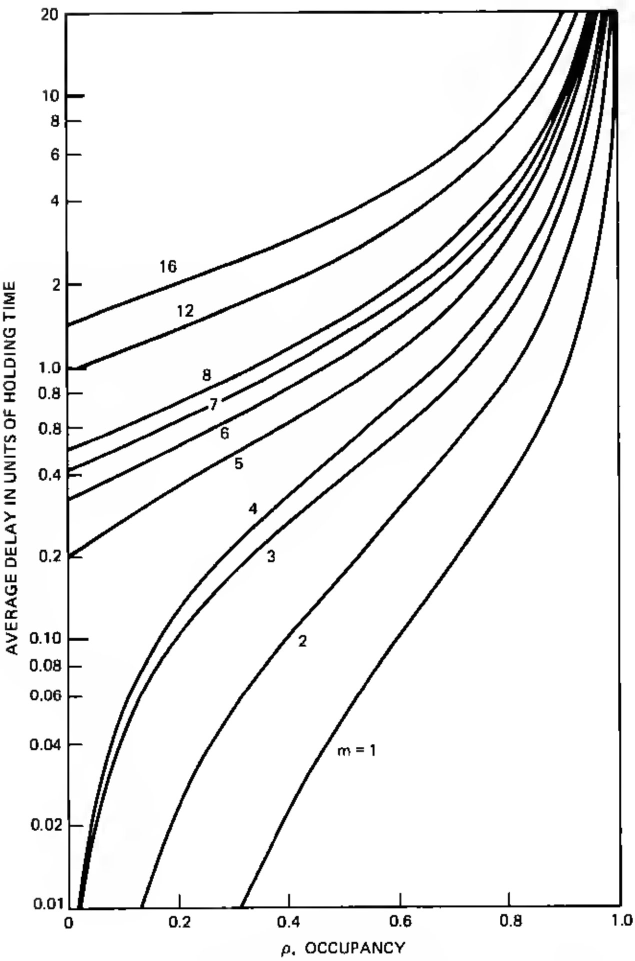

Fig. 1—Average delay in an M/D/4 queue when arrivals occur in batches of size m.

collection of single-server systems with constant holding time and j-phased Erlangian input of mean offered load μ_1/j. This can be seen by imagining that the sets of m servers required to serve the arriving batches are chosen in cyclic order.

Figure 2 shows the probability of service-delay (the probability that the service of an arriving batcb is delayed) as a function of the occupancy, for various values of the batch size m and $c = 4$. From eq. (12) we obtained and used the following expression for the service-

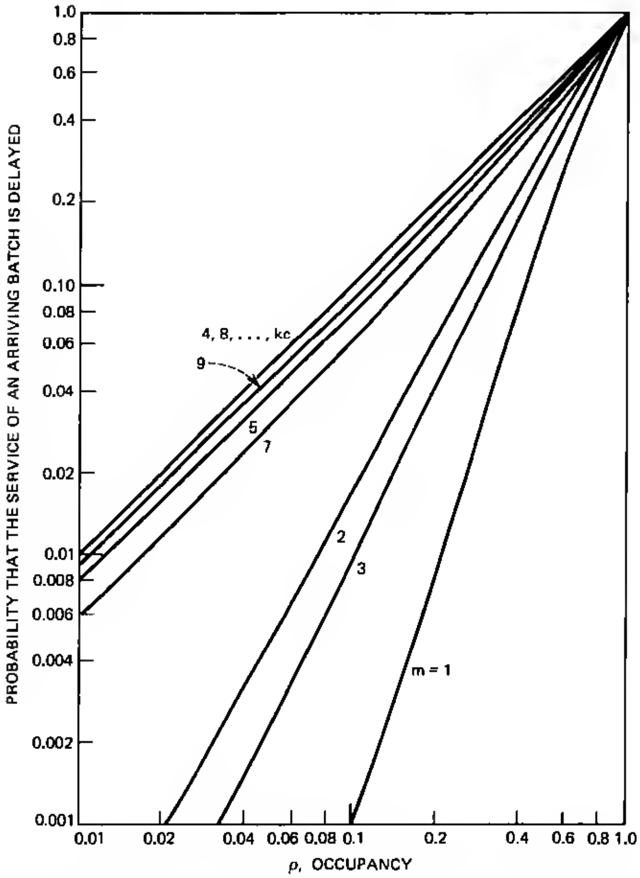

Fig. 2—Probability of service-delay in an M/D/4 queue when arrivals occur in batches of size m.

delay probability:

$$1 - \exp\left\{ - \sum_{n=1}^{\infty} \frac{1}{n} \left[1 - \sum_{j=0}^{nc-1} \pi_j(n) \right] \right\}.$$

Phenomena similar to tbose observed in Fig. 1 exist bere also. For example, if the batch size is an integer multiple of c, say $m = kc$, tben the probability of service-delay is the same as the probability of delay in a one-server system with single Poisson arrivals, i.e., it is simply ρ.

On the other hand, if $c = jm$, then the service-delay probability is the same as would be found in a system with single Poissonian arrivals of intensity λ_m and j constant-holding-time servers.

VII. ACKNOWLEDGMENT

I wish to thank A. Descloux of Bell Laboratories for suggesting the problem and for helpful discussions.

REFERENCES

1. Brockmeyer, E., Halstrom, M. L., and Jensen, Arne, *The Life and Works of A. K. Erlang*, Copenhagen: The Copenhagen Telephone Company, 1948.
2. Pollaczek, F., "Über eine Aufgabe der Wahrscheinlichkeitstheorie I," Mathematische Zeitschrift, *32*, 1930, pp. 64–100.
3. Pollaczek, F., "Über eine Aufgabe der Wahrscheinlichkeitstheorie II," Mathematische Zeitschrift, *32*, 1930, pp. 729–750.
4. Crommelin, C. D., "Delay Probability Formulae When the Holding Times Are Constant," P.O. Elec. Engrs. J., *25*, 1932, pp. 41–50.
5. Crommelin, C. D., "Delay Probability Formulae," P.O. Elec. Engrs. J., *26*, 1934, pp. 266–274.
6. Descloux, A., "Delay Systems With Random Input and Constant Holding Time," unpublished work.
7. Feller, W., *An Introduction to Probability Theory and Its Applications*, *1*, 2nd Edition, New York: Wiley, 1957.
8. Galliher, H. P., Morse, P. M., and Simond, M., "Dynamics of Two Classes of Continuous-Review Inventory Systems," Opns. Res., *7*, 1959, p. 362.
9. Ahlfors, L. V., *Complex Analysis*, 2nd Edition, New York: McGraw-Hill, 1966.